1. Find the mean and population standard deviation of the following set of data which represents the annual number of tornados in the U.S. for the years 1956 to 1975. Note that $\sum x = 14600$ and $\sum x^2 = 11221262$.

\[
\begin{array}{cccccccccccc}
504 & 856 & 564 & 604 & 616 & 697 & 657 & 464 & 704 & 906 \\
585 & 926 & 660 & 608 & 653 & 888 & 741 & 1102 & 947 & 918 \\
\end{array}
\]

**ANS:** First, the mean is $\mu = \frac{\sum x}{n} = \frac{14600}{20} = 730$.

For the standard deviation, we first compute

\[
SS_x = \sum x^2 - \left(\frac{\sum x}{n}\right)^2 = 11221262 - \left(\frac{14600}{20}\right)^2 = 563262.
\]

Therefore, the population standard deviation is

\[
\sigma = \sqrt{\frac{SS_x}{N}} = \sqrt{\frac{563262}{20}} \approx 167.819.
\]

2. Use Chebyshev’s theorem to find the interval centered about the mean for the annual number of tornados in which you would expect at least 75% of the years to fall.

**ANS:** Chebyshev’s theorem says that at least $(1 - \frac{1}{k^2}) \cdot 100\%$ of the data is within $k$ standard deviations of the mean. Using $k = 2$, this says at least $(1 - \frac{1}{2^2}) \cdot 100\% = 75\%$ of data is within 2 standard deviations of the mean. So we need to compute the interval that includes all numbers within 2 standard deviations of the mean. That is

\[
(\mu - 2\sigma, \mu + 2\sigma) = (730 - 2 \cdot 167.819, 730 + 2 \cdot 167.819) \approx (394, 1066).
\]

3. If the data in Question 1 is only considered a sample, compute the sample standard deviation.

**ANS:** Use the formula $s = \sqrt{\frac{SS_x}{n-1}}$ where $SS_x = 563262$ as computed above, and $n = 20$. Whence,

\[
s = \sqrt{\frac{563262}{19}} \approx 172.178.
\]
4. Consider the following grouped frequency distribution. Estimate the sample mean, sample standard deviation and coefficient of variation.

<table>
<thead>
<tr>
<th>x</th>
<th>12 - 14</th>
<th>15 - 17</th>
<th>18 - 20</th>
<th>21 - 23</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>2</td>
<td>4</td>
<td>11</td>
<td>3</td>
</tr>
</tbody>
</table>

**ANS:** We use the formulas $\bar{x} \approx \frac{\sum x \cdot f}{n}$ and $SS_x \approx \sum x^2 \cdot f - \frac{(\sum x \cdot f)^2}{n}$ where the midpoints and frequencies of each class are used for $x$ and $f$ respectively. Now,

\[
\sum x \cdot f = 13 \cdot 2 + 16 \cdot 4 + 19 \cdot 11 + 22 \cdot 3 = 365
\]

\[
\sum x^2 \cdot f = 13^2 \cdot 2 + 16^2 \cdot 4 + 19^2 \cdot 11 + 22^2 \cdot 3 = 6785
\]

Also, $n = \sum f = 2 + 4 + 11 + 3 = 20$. Therefore, $\bar{x} \approx \frac{365}{20} = 18.25$ and

\[
SS_x \approx 6785 - \frac{365^2}{20} = 123.75
\]

and so $s \approx \sqrt{\frac{123.75}{19}} \approx 2.55$. Finally, the coefficient of variation is

\[
\frac{s}{\bar{x}} \cdot 100% \approx 13.98%.
\]